

## ASSESSMENT OF GERMPLASM USING MULTIVARIATE ANALYSIS FOR GRAIN YIELD AND QUALITY TRAITS IN SPRING WHEAT

SYED BILAL HUSSAIN<sup>\*1</sup>, MUHAMMAD ATIF WAHID<sup>1</sup>, MUHAMMAD ZUBAIR<sup>2</sup>,  
MUHAMMAD BABAR<sup>3</sup> AND KASHIF WAHID<sup>1</sup>

<sup>1</sup>Department of Plant Breeding and Genetics, Bahauddin Zakariya University, Multan, Pakistan.

<sup>2</sup>Department of Forestry & Range Management, Bahauddin Zakariya University, Multan, Pakistan

<sup>3</sup>Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University, Multan, Pakistan

\*Corresponding author email: drsyedbilal@hotmail.com

### Abstract

Present study was designed to find diversity patterns among fourteen lines of hexaploid wheat (*Triticum aestivum* L.). Nine quantitative traits were determined phenotypically. Moreover the grain properties were also determined by laboratory procedures. Cluster analysis categorized cultivars into four groups. Based on Principal Component Analysis (PCA), the first seven components explained over 88% of genetic variation. Cluster analysis based on PCA using the first seven principal components indicated four separate groups of genotypes, with the maximum genetic distance observed between the genotypes in each cluster. Grain diameter, grain weight, gluten content, protein content, fiber, fat and starch contributed towards significant principal components (PCs). The genotypes 88146, Fareed-06, 6317 and 88132 were concluded as more diverse parents. Diverse parents from various clusters are helpful in planning and broadening the breeding programme by planning the crosses and increased use of heterosis and genetic diversity especially for grain quality in Pakistan.

### Introduction

Being staple food, wheat is principal source of carbohydrates in human diet. In Pakistan, wheat is grown as the major staple food for the people. Average annual production is almost equivalent to 24.032 million tons on 9.046 million hectares as renowned by (Hussain *et al.*, 2011); that comprised almost 34% of the total cultivated area in Pakistan. Increase in population demands an ever increasing demand for higher yield of this staple crop. In fact the current cultivars in wheat do not exhibit a lot of genetic diversity making it vulnerable to various biotic stresses. Economic value of bread wheat can be determined the primary and vital traits including grain yield with grain protein contents for its better end use quality (Oury & Godin, 2007). Therefore, it is necessary to utilize sources of new diversity in breeding. Researchers are exploiting different species or accessions with in genus Triticeae over the last 3 to 4 decades in order to broaden the wheat genetic base through enrichment of alleles with the aid of intergeneric crosses, intra-specific and interspecific (Mujeeb-Kazi *et al.*, 2008; Mujeeb-Kazi *et al.*, 2013; Ogbonnaya *et al.*, 2013).

Genetic diversity aids in identification of suitable parents which is an essential step in breeding of high yielded genotypes for future use. A thorough knowledge regarding to the existing genetic variability is required for the development of desirable traits wheat hybrid (Kahrizi *et al.*, 2010). Wheat yield changes widely as a result of its interaction with environment due to its complex inheritance and product of a number of contributing factors (Sajjad *et al.*, 2011). (Muhammad *et al.*, 2005) described positive and significant association between protein and gluten contents at genotypic level. Genetic diversity and numerical taxonomic techniques can be estimated by quantitative traits using principal component and cluster analysis to measure the pattern of genetic diversity in germplasm of different crops, as in black gram (Ghafoor *et al.*, 2001), chickpea (Naghavi & Jahansouz, 2005) and lentil (Sultana *et al.*, 2006). For

selection of genetically distance parents, several genetic diversity studies have been conducted among different crop species on the basis of qualitative and quantitative traits diversified parents in a hybridization programme. (Shekhawat *et al.*, 2001; Arega *et al.*, 2007; Haydar *et al.*, 2007; Ahmadizadeh *et al.*, 2011 & Daniel *et al.*, 2011).

To harness friable genetic variation in breeding material, it is worthwhile to trace the total variation into its components. Thus the present study was undertaken realizing the importance and need for such a comparative study in hexaploid wheat. Especially to study magnitude of genetic diversity using multivariate techniques on the basis of agronomic characters and quality parameter to identify superior accessions, choice of parents and to adapt an efficient breeding programme for variety development in country.

### Materials and Methods

A collection of 14 advanced lines including 3 standard check varieties of bread wheat (*T. aestivum* L.) were collected from Regional Agricultural Research Institute, Bahawalpur, Pakistan (Table 1). The germplasm was evaluated during crop season 2010-2011 at the Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences and Technology, Bahauddin Zakariya University, Multan. Randomized Complete Block Design with three replications was allocated. Each genotype was sown by single seed dibbler technique in 4 rows 5 meter long; keeping distance of 25 cm between rows and 15 cm within plants for each genotype. All the agronomic practices were applied uniformly to raise the crop. Data was recorded at an accurate time. Ten plants were randomly selected and tagged from central two rows to measure the following traits separately for each replication. Seven quantitative traits including plant height (PH), numbers of spikes per plant (S/P), spike length (SL), number of spikelets per spike (Sp/S), number of grains per spike (G/S), biomass of spike (BS), grain yield per plot (GY/P) were measured for each replication.

**Table 1. List of wheat genotypes used in experiments.**

Genotypes	Varietal code	Genotypes	Varietal code
Mairaj-08	V1	99199	V8
88131	V2	88146	V9
88132	V3	88148	V10
99192	V4	6317	V11
88106	V5	32862	V12
88124	V6	76346	V13
Fareed-06	V7	Sehar-06	V14

**Table 2. Mean squares of the morphological traits for wheat (*T. aestivum* L.).**

SOV	Df	Plant height (cm)	Spike length (cm)	Spikes/plant	Biomass of spike (g)	Spikelets/spike	Grains/spike	Yield/plot
Replication	2	57.71	0.65	0.74	0.042	1.36	1.023	0.052
Variety	13	62.83*	0.72*	0.40 <sup>ns</sup>	0.095*	1.51*	40.33**	0.035*
Error	26	25.20	0.32	0.28	0.038	0.61	4.357	0.016
Total	41							

\* = Significant, \*\* = Highly significant and ns = Non-significant

To assess the quality parameters a second experiment was performed at Cereals Laboratory, Ayub Agricultural Research Institute, Faisalabad for measuring grain physical characteristics such as thousand kernels weight, Single Kernel Characterization System (SKCS) (kernel weight, kernel diameter, kernel moisture, hardness index, grain length and grain width) Grain nutritional profile (protein content, starch content, gluten content, zeleny sedimentation rate) were measured by Kernelyzer (Omeg Analyzer G model) and ash content, crude fat and fiber were also measured.

The recorded data for all attributes was statistically analyzed using multivariate analysis such as Principal Component Analysis and Cluster analysis were performed for determination of genetic diversity based on morphological and quality attributes. Principal component analysis was based on correlation matrix. Statistically significant principal components were selected using Eigen significant criteria (Kaiser *et al.*, 1960), where as cluster analysis was applied according to (Ranjbar *et al.*, 2007 & Chaparzadeh *et al.*, 2008).

## Results and Discussion

Analysis of variance (ANOVA) represented that genotypes used for present study had significant variation for most of the traits. Spikes per plant (Table 2) and grain diameter (Table 3) showed non-significant variation among the genotypes for the year whereas grains per spike (Table 2), moisture (Table 3), protein content, fat and fiber (Table 4) were highly significant. Principal Component analysis is a classical technique used for data analysis, compression and visualization of features of data set (Jolliffe *et al.*, 1986) and (Bishop, 2006). Actually it reflects the determination of major contributor to the total variation at each axis of differentiation.

Eigen values helped in recognizing the detaining factor and their sum is approximately equal to total number of variables. Eigen values for 13 principal components have been shown in the scree plot (Fig. 1). The data revealed that

out of thirteen, seven PCs exhibited Eigen value (also called latent root) greater than one (significant) contributing 88.09% of variation (Tables 5 and 6). The rest six PCs explained trivial (non-significant) amount of variation, and were not worth interpreting. Among 14 accessions of wheat, the contributions by the first four PCs showed 65.76% of variability and were due importance for further explanation (Table 4). The PC<sub>1</sub> has 23.47%, PC<sub>2</sub> showed 19.20% and PC<sub>3</sub> exhibited 12.10% variability among the genotypes for the traits under study. In fact, with this method, 12 variables were reduced to seven. Gluten, protein, grain weight, grain diameter, starch, fat and fiber were noted as the characteristics for variability (Fig. 1). The principle developed by (Johnson and Wichern, 1988) was used to determine the importance of a trait coefficient for each significant principal component.

The first principal component was highly related to the gluten, protein, grains weight and grain diameter. This implies that PC<sub>1</sub> is a weighted average of these four traits. Whereas in second principal component, starch, fat and fiber are the traits of significant importance. The third principal component exhibited positive effect for moisture, grain width, zeleny sedimentation rate and 1000-grain weight (Table 7).

A principle component biplot showed that variables are super imposed on a plot as vectors; relative length of vector represents the relative proportion of the variability in each variable represented. One motivating explanation of biplot is that the cosine of the angle between the vectors of two indices estimates the correlation coefficient between them. As the biplot does not give explanation all of the deviations in a data set so cosine of the angles does not precisely interpret into correlation coefficients. However, the angles are instructive enough to draw a picture about the correlation among the different indices (Yan & Kang, 2003). In PC<sub>1</sub> and PC<sub>2</sub> together gluten, protein, grain weight and diameter are well represented in the plot but hardness index, biomass of spike, grains per spike and 1000-grains weight have minimum difference in PC<sub>1</sub> and PC<sub>2</sub> (Fig. 2).

**Table 3. Mean squares of the Kernel traits for wheat (*T. aestivum* L.).**

SOV	df	1000-kernel weight (g)	Hardness index %	Diameter (mm)	Grain weight (mg)	Grain length (mm)	Grain width (mm)	Moisture %
Variety	13	14.68*	45.32*	0.00986 <sup>ns</sup>	6.29*	0.35*	0.05*	0.62**
Error	28	6.47	20.83	0.00797	2.61	0.15	0.02	0.13
Total	41							

\* = Significant, \*\* Highly significant and ns = Non-significant

**Table 4. Mean squares of the Nutritional profile for wheat (*T. aestivum* L.).**

SOV	df	Protein	Starch	Gluten	Zeleny	Ash	Fat	Fiber
Variety	13	1.43**	1.30*	10.55*	147.50*	0.35**	0.060**	0.09146**
Error	28	0,38	0.55	4.90	56.71	0.008	0.005	0.01041
Total	41							

\* = Significant, \*\* = Highly significant and ns = Non-significant

**Table 5. Mean, minimum and maximum values of grain yield traits for the year 2010-11.**

Characters	Plant height (cm)	Spike length (cm)	Spikes/plant	Biomass of spike (g)	Spikelets/spike	Grains/spike	Yield/plot
Mean	91.09	8.91	5.38	2.23	16.35	34.90	1.70
Range	83.5-98.7	8.3-10.3	5.0-5.7	1.9-2.57	15.3-17.6	27-43.6	1.55-1.90

\* = Significant, \*\* = Highly significant and ns = Non-significant

**Table 6. Eigen values of correlation matrix and related statistics for first seven PC's active variables only.**

PCA	Eigen value	% Total - variance	Cumulative – Eigen value	Cumulative - %
1	4.927789	23.46566	4.92779	23.4657
2	4.032677	19.20322	8.96047	42.6689
3	2.538137	12.08636	11.49860	54.7552
4	2.311460	11.00695	13.81006	65.7622
5	1.901104	9.05287	15.71117	74.8151
6	1.590544	7.57402	17.30171	82.3891
7	1.197682	5.70325	18.49939	88.0923

**Table 7. Factor coordinates of the characters, based on correlations.**

Characteristics	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Plant height (cm)	0.7919	0.2644	0.1897	0.1215	0.2960	0.1294	0.1875
Spike length (cm)	0.3158	0.2779	0.2405	0.3672	0.6840	0.0093	0.0966
Tillers	0.2979	0.4605	0.1271	0.5767	0.0790	0.2648	0.2727
Biomass of spike (g)	0.4179	0.4786	0.5929	0.0163	0.1292	0.0162	0.2698
Spikelets /spike	0.4859	0.2576	0.6180	0.0863	0.2014	0.1277	0.1789
Grains /spike	0.2970	0.3756	0.7298	0.2095	0.2635	0.0419	0.1950
Yield /plot (kg)	0.4546	0.3510	0.3512	0.2417	0.5710	0.1580	0.1523
1000 Grains weight (g)	0.2599	0.1982	0.3069	0.1320	0.1736	0.6167	0.5507
Hardness index (%)	0.4164	0.4626	0.4049	0.0611	0.3496	0.4696	0.0833
Diameter (mm)	0.7032	0.1705	0.2190	0.4491	0.3791	0.0524	0.0998
Grain weight (mg)	0.7343	0.1491	0.0244	0.5136	0.3829	0.0882	0.0030
Grain length (mm)	0.1206	0.2321	0.3958	0.6566	0.1420	0.1635	0.1735
Grain width (mm)	0.3029	0.0874	0.4062	0.5956	0.0838	0.4614	0.0091
Protein (%)	0.7701	0.5591	0.0211	0.0095	0.2398	0.0572	0.0877
Moisture (%)	0.6670	0.2247	0.4568	0.0861	0.3940	0.0541	0.0636
Starch (%)	0.0772	0.7062	0.1837	0.4754	0.1409	0.3335	0.1345
Gluten (%)	0.7720	0.5423	0.0196	0.0162	0.1942	0.0028	0.2017
Zeleny (%)	0.2668	0.8037	0.3094	0.0760	0.0290	0.1987	0.0841
Ash (%)	0.2120	0.5573	0.0834	0.3364	0.2920	0.3290	0.4745
Fat (%)	0.1454	0.5756	0.1926	0.1053	0.0487	0.5862	0.4408



Fig. 1. Scree plot between Eigen values and number of principal components.

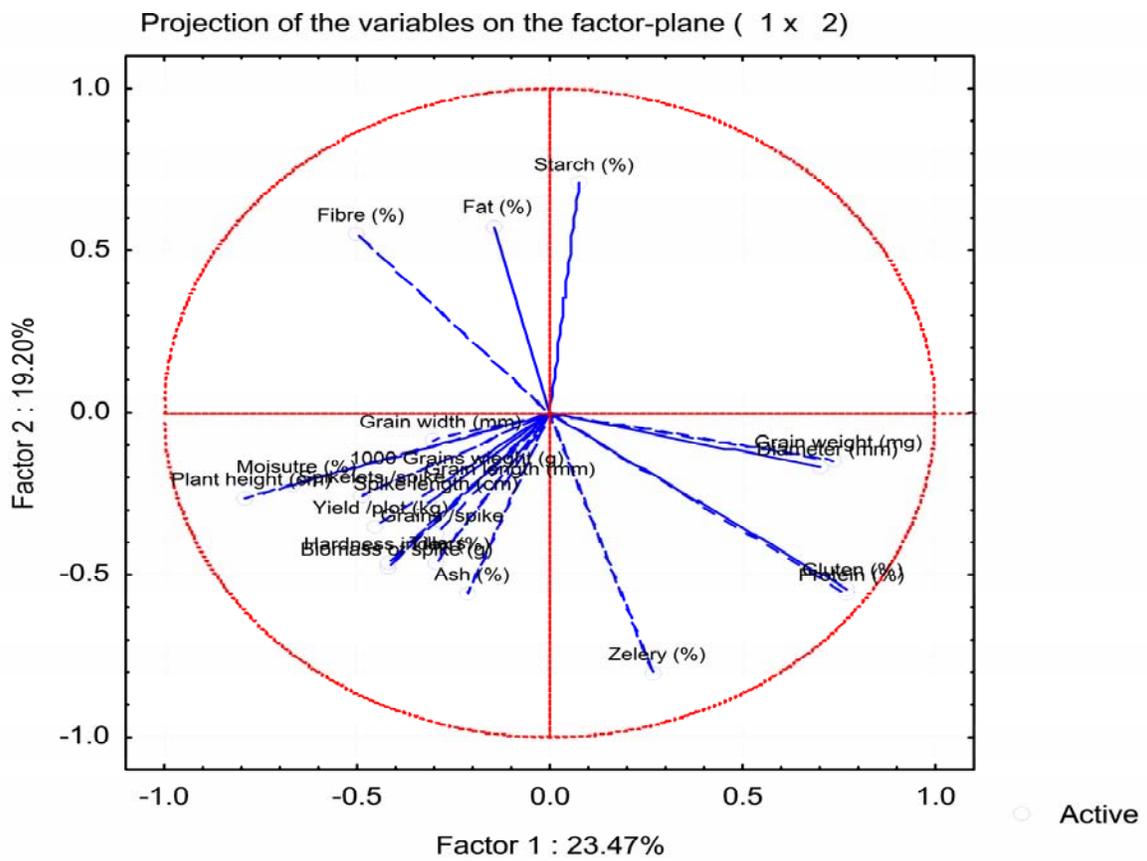


Fig. 2. Principal component biplot of 14 wheat germplasm lines.

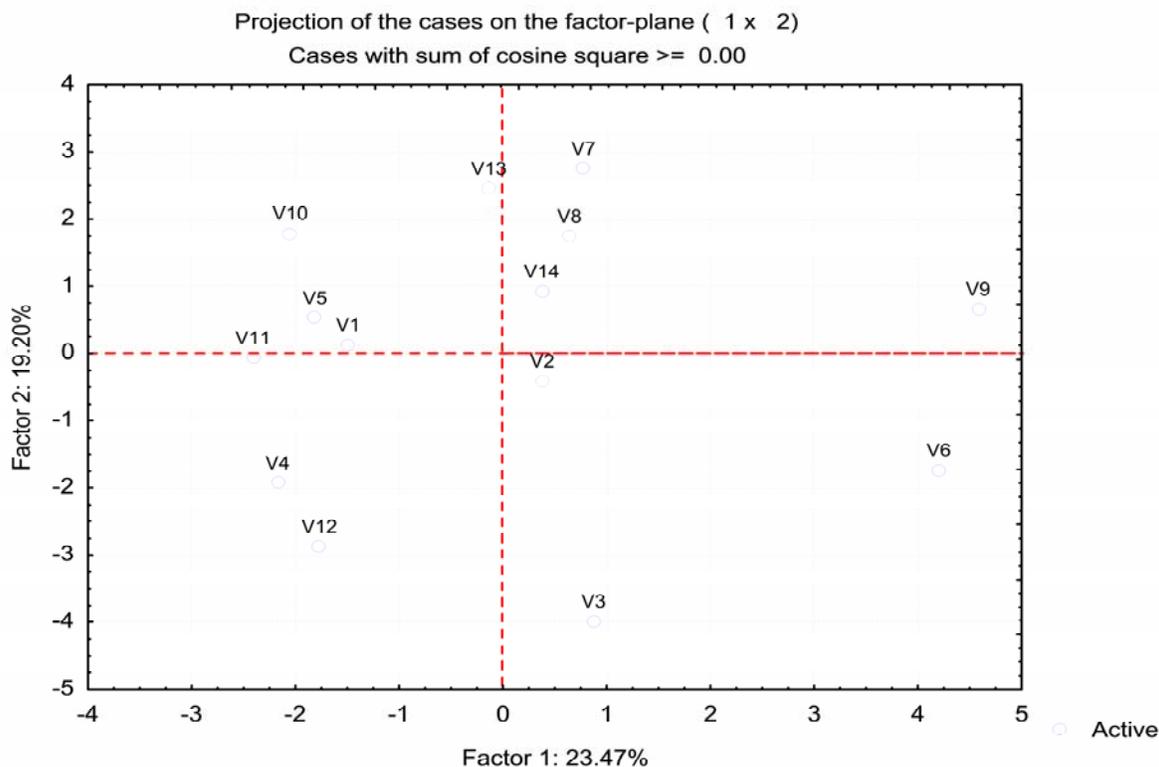


Fig. 3. Two dimensional ordinations of 14 germplasm lines of wheat on principal component axis I and II.

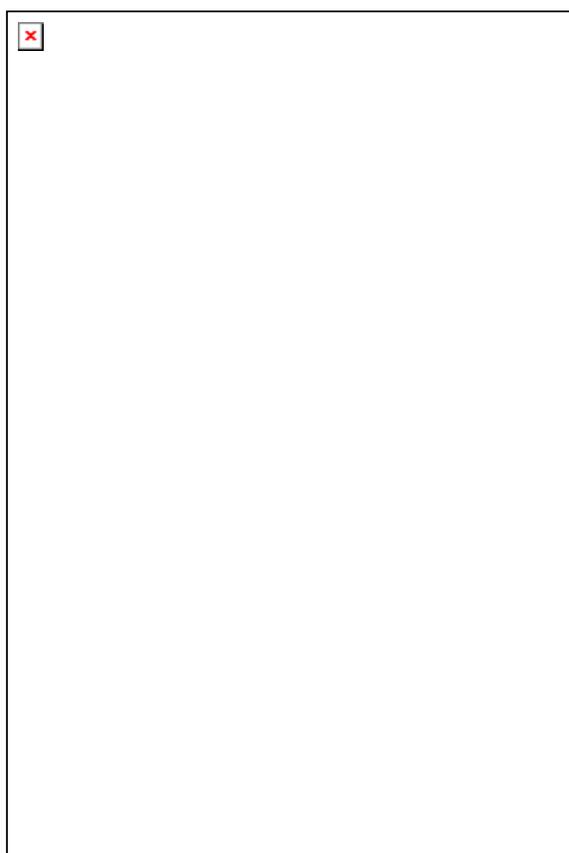


Fig. 4. Dendrogram resulting from cluster analysis of 14 wheat genotypes on the basis of all traits.

A principal component scatter plot of the wheat accessions depicts that the accessions that are close together are perceived as being similar when rated on the twenty one variables (Fig.3). Accessions which are further apart are more different. Thus most of accessions are close to each other while V3 (88132), V6 (88124), V7 (Fareed-06), V9 (88146), V11 (6317), and V13 (76346) are rather separated from each other on both PC<sub>1</sub> and PC<sub>2</sub>. This population structure aids in identification of diverse groups of parents from segregating population. The results agreed well with (Sajjad *et al.*, 2011) that genotypes of PC<sub>1</sub> and PC<sub>2</sub> showed population structure for two years data.

The dendrogram deliberate the relative magnitude of resemblance among the genotypes as well as the clusters. On the basis of first seven principal components of PCA Cluster analysis indicated three separate groups of genotypes, with the maximum genetic distance observed between and genotypes. According to dendrogram obtained (Fig. 4) cluster analysis grouped V1 (Mairaj-08) and V4 (99192) in cluster I. Cluster II carried, V8 (99199), V14 (Sehar-06), V10 (88148), V11 (6317), V13 (76346) and V7 (Fareed-06). While cluster III carried V2 (88131), V5 (88106), V3 (88132), V12 (32862), V6 (88124) and V9 (881446).

Cluster I carried the genotypes that indicated the lowest genetic difference among them. Cluster III carried the genotypes which had large genetic diversity. On the basis of comparison among the genotypes grouped in the same cluster it was clear that the genotypes grouped in cluster I exhibited least differences among themselves (Fig. 4).

There is significant genetic variability among tested genotypes indicating the presence of tremendous opportunity for improving through distant hybridization by

crossing genotypes with high genetic distance. PCA scores revealed greatest diversity between 88146 and Fareed-06 while lowest between 6317 and 88312 similar with the findings of (Singh *et al.*, 2014). The information obtained from this study can be used to plan crosses and maximized the use of genetic diversity and expression of quality traits as similar explained by (Ahmad *et al.*, 2013). The number of spikes per plant for 2010-2011 for current genotypes (5-5.7) was lower than found in Indian germplasm, (Kumar *et al.*, 2009), West Bengal germplasm, (Kotal *et al.*, 2010), Pakistani germplasm and CIMMYT (Gulnaz *et al.*, 2011). The results are similar to that of determined by using principal component analysis (Maqbool *et al.*, 2010).

Various Scientific studies at CIMMYT revealed that Synthetic wheat varieties carry tremendous bread-making quality can definitely be bred if the common bread wheat parents in the cross have excellent bread quality traits. HMW and LMW glutenin profiles of the synthetic hexaploids (Rasheed *et al.*, 2012) can be utilized to identify promising crosses and best quality lines in their progeny. Nelson *et al.*, (2006) described that some lines from International Triticeae Mapping Initiative population showed better quality values as compared to their parental lines.

## Conclusion

The genetic base of the wheat breeding programme can be increased using the diverse derivation of the genotypes present in this collection of germplasm and will contribute to new alleles for both grain yield and grain quality.

## References

- Ahmad, A., A. Muhammad, A. Maria Mastarengel, P. De Vita, A. Gul. Kazi and A. Mujeeb-Kazi. 2013. Comparative assessment of glutenin composition and its Relationship with grain quality traits in bread wheat and synthetic derivatives. *Pak. J. Bot.*, 45(SI): 289-296.
- Ahmadzadeh, M., H. Shahbazi, M. Valizadeh and M. Zaefizadeh. 2011. Genetic diversity of durum wheat landraces using multivariate analysis under normal irrigation and drought stress conditions. *Afr. J. Agric. Res.*, 6(10): 2294-2302.
- Bishop, C. 2006. Pattern Recognition and Machine Learning. Springer, Heidelberg.
- Chaparzadeh, N.O., A. Sofalian, A. Javanmard, M.S. Hejazi and L. Zarandi. 2008. Study of glutenin subunits in some wheat landraces from northwest of Iran by SDS-PAGE technique. *Int. J. agric. Biol.*, 10(1): 101-104.
- Daniel, H., M. Mebrahtom and G. Tsige. 2011. Genetic divergence analysis on some bread wheat genotypes grown in Ethiopia. *J. Central Eur. Agric.*, 12(2): 344-352.
- Ghafoor, A., A. Sharif, Z. Ahmad, M.A. Zahid and M.A. Rabbani. 2001. Genetic diversity in Blackgram (*Vigna mungo* L.) Hepper). *Field Crops Res.*, 69: 183-190.
- Gulnaz, S., M. Sajjad, I. Khaliq, A.S. Khan and S.H. Khan. 2011. Relationship among coleoptile length, plant height and tillering capacity for developing improved wheat varieties. *Int. J. Agric. Biol.*, 13: 130-133.
- Haydar, A., M.B. Ahmed, M.M. Hannan, M.A. Razvy, M.A. Mandal, M. Salahin, R. Karim and M. Hossain. 2007. Analysis of genetic diversity in some potato varieties grown in Bangladesh. *Middle-East J. Sci. Res.*, 2(3-4): 143-145.
- Hussain, M., N. Ahmad, F. Mohammad, M.A. Khan and S. Ajmal. 2011. Wheat breeding for high yield potential and durable resistance against yellow rust. *Pak. J. Phytopathol.*, 23: 56-61.
- Johnson, R.A and D.W. Wichern. 1988. Applied multivariate statistical analysis. Prentice- Hall, Englewood Cliffs, NJ.
- Jolliffe, I. 1986. Principal Component Analysis. Springer, Heidelberg.
- Kahrizi, D., M. Maniee, R. Mohammadi and K. Cheghamirza. 2010b. Estimation of genetic parameters related to morpho-agronomic traits of Durum wheat (*Triticum turgidum* var. durum). *Biharean Biologist.*, 4(2): 93-97.
- Kaiser, H.F. 1960. The application of electronic computers to factor analysis. *Edu. Psych. Measurement.*, 20: 141-151.
- Kotal, B.D., A. Das and B.K. Choudhary. 2010. Genetic variability and association of characters in wheat (*T. aestivum* L.). *Asian J. Crop Sci.*, 2: 155-160.
- Kumar, B., G.M.R. Lal and A. Upadhyay. 2009. Genetic variability, diversity and association of quantitative traits with grain yield in bread wheat (*T. aestivum* L.). *Asian J. Agric. Sci.*, 1: 4-6.
- Muhammad, T., S. Haider, M. Amin., M.I. Khan and R. Zamir. 2005. Path coefficient and correlation studies of yield and yield associated traits in candidate bread wheat (*T. aestivum* L.) Lines. *Suran. J. Sci. Technol.*, 13: 175-180.
- Mujeeb-Kazi, A., A. Gul, M. Farooq, S. Rizwan and I. Ahmad. 2008. Rebirth of synthetic hexaploids with global implications for wheat improvement. *Aust. J. Agric. Res.*, 59: 391-398.
- Naghavi, M.R. and M.R. Jahansouz. 2005. Variation in the agronomic and morphological traits of Iranian chickpea accessions. *J. Integ. Plant Biol.*, 47(3): 375-379.
- Nelson, J.C., C. Andreescu, F. Breseghello, P.L. Finney, G. Daisy, D.G. Gualberto, C.J. Bergman, R.J. Pena, M.R. Perretant, P. Leroy, C.O. Qualset and M.E. Sorrells. 2006. Quantitative trait locus analysis of wheat quality traits. *Euphytica.*, 149: 145-159.
- Ogbonnaya, F., A. Mujeeb-Kazi, A. G. Kazi, E.L. Lagudah, S.S. Xu and D. Bonnett. 2013. Synthetic hexaploid in wheat improvement. In: (Ed.): Jules Janick. pp. 35-122. Plant Breeding Reviews. John Wiley & Sons Inc.
- Oury, F.X. and C. Godin. 2007. Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favorable genotypes? *Euphytica*, 157: 45-57.
- Ranjbar, M., M.R. Naghavi, A. Zali and M.J. Zaghaei. 2007. Multivariate analysis in morphological variation in accessions of *Aeglops crassa* from Iran. *Pak. J. Biol. Sci.*, 10(7): 1126-1129.
- Rasheed, A., T. Mahmood, A.G. Kazi, A. Ghafoor and A. Mujeeb-Kazi. 2012. Allelic variation and composition of HMW-GS in advanced lines derived from D-genome synthetic hexaploid / bread wheat (*Triticum aestivum* L.). *J. Crop Sci. Biotech.*, 15(1): 1-7.
- Sajjad, M., S.H. Khan and A.S. Khan. 2011. Exploitation of germplasm for grain yield improvement in spring wheat (*T. aestivum* L.). *Int. J. Agric. Biol.*, 13:695-700.
- Shekhawat, U.S., P. Vijay and D.L. Singhania. 2001. Genetic divergence in barley (*Hordeum vulgare* L.). *Indian J. Agric. Res.*, 35(2): 121-123.
- Singh, P., A.K. Singh, M. Sharma and S.K. Salgotra. 2014. Genetic divergence study in improved bread wheat varieties (*Triticum aestivum*). *African J. Agric Res.*, 9(4): 507-512.
- Sultana, T., A. Ghafoor and M. Ashraf. 2006. Geographic patterns of diversity of cultivated lentil germplasm collected from Pakistan, as assessed by seed protein assays. *Acta Biologica Cracoviensia, Series Botanica, Poland.*, 48(1): 77-84.
- Welch, R.M and R.D. Gram. 2004. Breeding for micronutrients in staple food crops from a human nutrition perspective. *J. Exp. Bot.*, 55: 353-64.
- Yan, W. and M.S. Kang. 2003. Biplot Analysis: A graphical tool for breeders, geneticists and agronomist, CRC Press, Boca Raton, FL. 313.